

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 3 年 1 1 月 1 3 日
Date of Application:

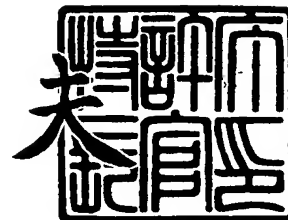
出 願 番 号 特 願 2 0 0 3 - 3 8 3 2 4 8
Application Number:
[ST. 10/C] : [J P 2 0 0 3 - 3 8 3 2 4 8]

出 願 人 株式会社日立製作所
Applicant(s):

2 0 0 3 年 1 2 月 1 0 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康



【書類名】 特許願
【整理番号】 K03010891A
【あて先】 特許庁長官殿
【国際特許分類】 G06F 12/00
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 レモアル ダミエン
【発明者】
 【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所
 システム開発研究所内
 【氏名】 峯 博史
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社 日立製作所
【代理人】
 【識別番号】 100075096
 【弁理士】
 【氏名又は名称】 作田 康夫
【選任した代理人】
 【識別番号】 100100310
 【弁理士】
 【氏名又は名称】 井上 学
【手数料の表示】
 【予納台帳番号】 013088
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

ホストコンピュータとメインストレージとを相互接続するデバイスアダプタであって、
前記デバイスアダプタはキャッシュストレージにアクセスするためのキャッシュストレージインタフェース装置、及び、ネットワークを介して前記メインストレージにアクセスするメインストレージインタフェース装置と、前記ホストコンピュータによるデータの出力要求を処理するキャッシュコントローラとを具備し、前記キャッシュコントローラは、前記メインストレージから取り出したデータを前記キャッシュストレージにキャッシュする処理部と、

前記キャッシュストレージの記憶領域の割当と使用状況を管理する処理部と、

前記メインストレージと前記キャッシュストレージとに関する負荷とパフォーマンスの情報を管理する処理部と、

前記要求されたデータがキャッシュされているか否かを判断する処理部と、

前記要求されたデータがキャッシュされている時に、前記要求データを取り出すストレージを、前記キャッシュストレージ、及び、メインストレージのいずれとするかを決定する処理部とを備える。

【請求項 2】

請求項 1 に記載のデバイスアダプタであって、

前記キャッシュする処理部は、

前記取り出したデータが、前記キャッシュストレージに既にキャッシュされている、或いは、前記キャッシュストレージインタフェース装置のデータ転送帯域幅が予め設定した閾値を超えている場合には、前記取り出したデータを、前記キャッシュストレージにキャッシュしない。

【請求項 3】

請求項 1 に記載のデバイスアダプタであって、

前記負荷の情報は、前記メインストレージインタフェース装置および／または前記キャッシュストレージインタフェース装置の現在のデータ転送速度であって、

前記要求データを取り出すストレージを決定する処理部は、前記現在のデータ転送速度を参照して決定する。

【請求項 4】

請求項 1 に記載のデバイスアダプタであって、

前記パフォーマンスの情報は、前記メインストレージインタフェース装置および／または前記キャッシュストレージインタフェース装置の平均要求処理時間であって、

前記要求データを取り出すストレージを決定する処理部は、前記平均要求処理時間を参照して決定する。

【請求項 5】

請求項 1 に記載のデバイスアダプタであって、

前記メインストレージインタフェース装置は、前記キャッシュストレージインタフェース装置とは異なるインタフェースを備える。

【請求項 6】

請求項 3 に記載のデバイスアダプタであって、

前記メインストレージインタフェース装置は、ファイバーチャネルインタフェースを備える。

【請求項 7】

請求項 3 に記載のデバイスアダプタであって、

前記キャッシュストレージインタフェース装置は、SCSI インタフェースを備える。

【請求項 8】

請求項 1 に記載のデバイスアダプタであって、

前記キャッシュコントローラは、さらに、前記キャッシュストレージの領域の割当て処理部を備える。

【請求項 9】

請求項 6 に記載のデバイスアダプタであって、
前記キャッシュストレージ領域割当て処理部は、前記キャッシュストレージの領域が利用できない場合に、置換すべきデータを選択して置換する。

【請求項 1 0】

請求項 1 に記載のデバイスアダプタであって、
前記要求されたデータがキャッシュされているか否かを判断する処理部は、前記判定にディスクセクタ番号を用いる。

【書類名】 明細書**【発明の名称】 統合キャッシュを備えた記憶装置アダプタ****【技術分野】****【0001】**

本発明は、ストレージエリアネットワークからのデータの取出しとキャッシングに関し、特に、OSのサポートや変更を行うことなくストレージからアクセスされるデータのキャッシングを可能にするデバイスアダプタに関する。

【背景技術】**【0002】**

SAN (Storage Area Network) を介してアクセスされる、RAID (安価なディスクの冗長アレイ) などのネットワークアクセス型ストレージシステムは、信頼性と記憶容量の拡張性の点から極めて有用であり、非常に高速なデータ入出力 (I/O) を可能とする。SANを実現するために使用されるファイバーチャネル (FC) ネットワークやIPネットワークによって、さらにはネイティブSCSIコマンドやSCSI/IPプロトコル、例えばiSCSIを用いて、そのようなネットワークアクセス型ストレージシステムに非常に柔軟にアクセスでき、その結果SANに接続された複数のストレージ装置に複数のホストコンピュータから同時に直接アクセスすることが可能となる。

【0003】

上述のネットワークアクセス型ストレージシステムにおいては、複数のホストコンピュータがSANに接続されたストレージ装置のデータに同時にアクセスすることによって入出力速度が制限されてしまい、システム全体のパフォーマンスを低下させる可能性がある。具体的には、ストレージ装置からホストコンピュータへのデータ転送速度は、SANを構成するネットワーク媒体の使用可能な物理的帯域幅によって制限される。

【0004】

これにより、ストレージ装置が高速にデータを出力できても、その性能を十分に発揮することができない。また、SANを構成するスイッチは、複数のポート間において、ホストコンピュータからストレージ装置へあるいはストレージ装置からホストコンピュータへ転送するデータを処理する。そのため、広い入出力の内部帯域幅を必要とし全ての入出力動作を効率よく処理する必要がある。

【0005】

ストレージシステムに接続されたネットワークからデータをアクセスするホストコンピュータにおいては、前述の問題によって、データ転送速度の低下と、入出力要求処理に必要な時間の増加という2つの問題が生じる。ホストコンピュータをマルチメディアストリーミングサーバとして用いる場合には、クライアントに配信するためにストレージシステムからデータを取り出す能力によってサーバのパフォーマンスが制限されるため、上述の問題は特に重大である。

【0006】

具体的には、データの取出速度が遅いと、ホストコンピュータの演算処理能力に関わらずサーバとしてのパフォーマンスは低下する。リアルタイム処理を必要とするマルチメディアデータの配信を考えると、入出力要求の処理時間が増えることにより、データ伝送遅延が生じるため、クライアント側に配信、復号、提供するマルチメディアの品質を劣化させる恐れがある。

【0007】

多くのストレージシステムではこれらのパフォーマンス問題に取り組んでおり、それらを解決するために異なる解決法が提示されている。

【0008】

まず、使用している記憶媒体 (例えば磁気ディスクやテープ) のアクセス遅延を避け、ストレージシステムの入出力処理の遅延を克服するために、多くのストレージシステムでは、高速なキャッシュメモリを実装し、解決を図っている。このキャッシュメモリによって、頻繁にアクセスするデータのキャッシングが可能となり、その結果使用されている低

速な記憶媒体へのアクセスが回避され、そうしたデータへの入出力要求の処理時間が削減される（例えば、特許文献1参照）。

【0009】

また、データがホストコンピュータの近くにキャッシュされるキャッシング方式は、SAN内でのデータ転送帯域幅の消費量が削減できるのでより良い解決法である。ホスト拡張バスかあるいはSANにアクセスするために使用するデバイスアダプタに接続されるキャッシュデバイスが、いくつか提案されている。

【0010】

たとえば、特許文献2記載のデバイスアダプタにおいては、キャッシュデバイスはホストコンピュータの拡張バス（I/Oバス）に直接接続される。この方法では、大量のデータをキャッシュするために、キャッシュ記憶媒体として非常に大容量のディスクを使用し、非常に柔軟なキャッシングを実現している。

【0011】

また、ストレージシステムに接続されSANからデータを取り出すために使用されるデバイス内でキャッシング作業を行う方法がある。そのデバイスを介してアクセスするストレージへのデータキャッシュの有用性は縮小するが、キャッシング作業中のホスト拡張バスの過負荷を取り除くことができる。

【0012】

そのようなデバイスの実施例として、特許文献3記載のデバイスアダプタでは、IPネットワークに接続されたストレージシステムから取り出されたデータをアクセス装置に接続されたディスクにキャッシュすることができる。従って、キャッシング作業は、ホスト拡張バスに過負荷を与えることなくデバイスアダプタ内で実施することができる。

【0013】

また、データキャッシュに接続されたホストコンピュータの例として、アクセスしたストレージシステムと使用しているキャッシュストレージ装置の最高データ転送速度と処理速度の違いを考慮しないキャッシング方式が提示されている。

【0014】

【特許文献1】特開第2001-051901号公報

【0015】

【特許文献2】米国特許第6463509号公報

【特許文献3】国際公開第03/017598号パンフレット

【発明の開示】

【発明が解決しようとする課題】

【0016】

しかし、特許文献1に記載の方法では、キャッシングはアクセス先のストレージ装置内で行われるので、SAN内のネットワーク帯域幅の使用を削減することはできない。従って、この方法では、SAN内のデータ転送パフォーマンスの低下を解決できない。

【0017】

さらに、通常キャッシュのサイズは大量のデータを保持するほど十分な大きさはなく、ストリーミングサーバによってアクセスされるマルチメディアデータファイルでは、複数のホストサーバが続けて大容量のデータファイルに同時にアクセスした場合、キャッシュ全体の効率は低下してしまう。

【0018】

特許文献2に記載の方法では、異なる種類のインタフェースを介して様々な記憶装置から取り出されたデータは、ホストオペレーティングシステム内のデータ転送経路やデバイスドライバを変更することによって、それらのキャッシュデバイスにキャッシュすることができる。

【0019】

この方法によって、高い柔軟性がもたらされるが、2つの重大な問題が生じる。1つは、データのキャッシングおよびキャッシュデバイスからの取出しを処理するためにはオペ

レーティングシステムからのサポートが必要となることである。利用対象のオペレーティングシステムは、デバイスドライバの階層構造をサポートする必要がある。

【0020】

もう1つの問題は、キャッシングプロセス用拡張バスの帯域幅の消費量が過剰になることである。使用するバスの種類に応じて、キャッシュされるデータは拡張バスを介して2回転送される可能性がある。1回は記憶データにアクセスするデバイスからホストメモリへ、もう1回はホストメモリからキャッシュデバイスへと転送される。このようなホストI/Oバスの過負荷によって、マルチメディアストリーミングなどの入出力が集中する用途でのパフォーマンスが低下する可能性がある。

【0021】

特許文献3に記載の方法では、デバイスは、使用しているキャッシュストレージの負荷と、キャッシュされていないデータを取り出すために使用されるインタフェースの負荷とを正確に測定するための帯域幅制御を行う機能を持たず、データの取り出し先はキャッシュ内にデータがあるか否かによって判断される。インタフェースの負荷状況を用いて動的にデータの取り出し先の決定を行う機能は、マルチメディアデータなどのリアルタイム処理を特徴とするデータの取出しの場合に必要である。

【0022】

特定の種類の入出力処理に対して使用中のキャッシュデバイスがストレージ装置に接続したSANより低いパフォーマンスを示した場合、この方式ではシステム全体のパフォーマンスが低下する可能性がある。得られた入出力要求処理時間（もしくは少なくともその概算）がキャッシュデバイスの入出力要求処理に必要な入出力要求処理時間より短い場合、キャッシュ済みのデータであっても、SANを介するストレージシステムから取り出す必要が生じる可能性がある。

【0023】

その典型的な例としては、小さなデータブロックのデータの取出しがある。小さなデータブロックの取出しは通常、ディスクストライピングの効果によりディスクアレイシステムでは速くなり、ディスクシークのオーバーヘッドが大きいシングルディスクシステムでは遅くなる。

【課題を解決するための手段】

【0024】

本発明は、ホストコンピュータがネットワークを介してメインストレージから取り出したデータを有効にかつ柔軟にキャッシュして、キャッシュストレージの柔軟な大きさと、効率的なキャッシングプロセスと、ホストコンピュータによる入出力要求の処理時間を削減するためのデータ取出し先の動的制御とを行う、ホストコンピュータとメインストレージとを接続するデバイスアダプタを提供する。

【0025】

本発明による、ホストコンピュータとメインストレージとを相互接続するデバイスアダプタは、キャッシュストレージにアクセスするためのキャッシュストレージインタフェース装置、及び、ネットワークを介して前記メインストレージにアクセスするメインストレージインタフェース装置と、前記ホストコンピュータによるデータの出力要求を処理するキャッシュコントローラとを備える。

【0026】

キャッシュコントローラは、メインストレージから取り出したデータを前記キャッシュストレージにキャッシュする処理部と、キャッシュストレージの記憶領域の割当と使用状況を管理する処理部と、メインストレージと前記キャッシュストレージとに関する負荷とパフォーマンスの情報を管理する処理部と、要求されたデータがキャッシュされているかを判断する処理部と、要求されたデータがキャッシュされている時に、要求データを取り出すストレージを、キャッシュストレージ、及び、メインストレージのいずれとするかを決定する処理部とを備える。

【0027】

さらに、上記デバイスアダプタにおいて、上記キャッシュする処理部は、取り出したデータが、キャッシュストレージに既にキャッシュされている、或いは、キャッシュストレージインタフェース装置のデータ転送帯域幅が予め設定した閾値を超えている場合には、取り出したデータを、前記キャッシュストレージにキャッシュしないことを特徴とする。

【0028】

また、上記負荷の情報とは、メインストレージインタフェース装置および／またはキャッシュストレージインタフェース装置の現在のデータ転送速度であって、要求データを取り出すストレージを決定する処理部は、上記現在のデータ転送速度を参照して決定することを特徴とする。

【0029】

また、上記パフォーマンスの情報とは、メインストレージインタフェース装置および／またはキャッシュストレージインタフェース装置の平均要求処理時間であって、要求データを取り出すストレージを決定する処理部は、上記平均要求処理時間を参照して決定することを特徴とする。

【発明の効果】

【0030】

本発明によれば、ホストコンピュータに接続されたデバイスに記憶装置にネットワークを介して接続されたストレージから取り出したデータを有効にかつ柔軟にキャッシュして、キャッシュストレージの柔軟な大きさや、システムに負荷をかけない効率的なキャッシングプロセスや、ホスト入出力要求の処理時間を削減するためのデータ取り出し先の動的制御が可能になる。

【発明を実施するための最良の形態】

【0031】

図1には、一般的なホストコンピュータ(100)に接続された、本発明のデバイスアダプタ(200)が示されている。ホストコンピュータ(100)が備える一つ以上のCPU(中央処理装置)(110)は、コントローラ／ブリッジ(130)によって管理されるメモリバスを介してメインメモリ(120)に記憶されたプログラムをアクセスし実行する。メインメモリ(120)が記憶するプログラムは、デバイスドライバプログラムによるデバイス制御などのアプリケーションに様々なサービスを提供するホストオペレーティングシステム(OS)とアプリケーションプログラムを含む。

【0032】

デバイスアダプタ(200)からメインメモリ(120)へのデータ転送は拡張バス(140)を介して実施される。拡張バス(140)は、図示したような単一のバスではなく複数のバスから構成されても良い。あるいは拡張バスを備えず、デバイスアダプタ(200)がコントローラ／ブリッジ(130)によって管理されるメモリバスに直接接続されてもよい。

【0033】

本実施例のデバイスアダプタ装置(200)は、拡張バスインタフェース装置(210)を用いて、拡張バス(140)を介してホストコンピュータ(100)に接続される。

【0034】

デバイスアダプタ装置(200)は、キャッシュコントローラ(220)と、メインストレージインタフェース装置(240)と、キャッシュストレージインタフェース装置(250)とを備える。キャッシュコントローラ(220)は、メインストレージインタフェース装置(240)と、キャッシュストレージインタフェース装置(250)と、キャッシュディレクトリ(222)の使用量と、を監視するストレージインタフェースモニタ(221)を備え、ローカルキャッシュであるキャッシュストレージ(400)上の記憶領域割当てを管理する。

【0035】

デバイスアダプタ装置(200)の各構成要素は、ハードウェアまたはプロセッサがプログラムを実行することにより具現化され、以下に説明する処理フローは、デバイスアダ

プタが備えるプロセッサが、プログラムを実行するかあるいはハードウェアを制御することによって達成される。

【0036】

メインストレージ (500) は、コントローラ (241) を含むメインストレージインタフェース装置 (240) とネットワーク 300 を介してアクセスされる。キャッシュストレージ (400) は、コントローラ (251) を含むキャッシュストレージインタフェース装置 (250) を用いてアクセスされる。

【0037】

メインストレージインタフェース装置 (240) とキャッシュストレージインタフェース装置 (250) は、デバイス統合バス (230) に接続され、拡張バスインタフェース装置 (210) とキャッシュコントローラ (220) を、介して拡張バス (140) とのデータ転送を実施する。

【0038】

デバイスアダプタ (200) は、ホストコンピュータ (100) からは、単一インタフェース、即ちメインストレージインタフェース装置 (240) として動作する。すなわち、キャッシュコントローラ (220) とキャッシュストレージ (400) は隠され、キャッシュストレージ制御とキャッシング処理は、キャッシュコントローラ (220) が実装されたデバイスアダプタ (200) 内でホストコンピュータ (100) に意識されずに実行される。

【0039】

図 2 は、デバイスアダプタ (200) による、ホストコンピュータ (100) によって発せられた読出し入出力要求の処理フローを示す。

【0040】

キャッシュコントローラ (220) が読出し要求を受けると (s100)、キャッシュディレクトリ (222) を使用してデータがキャッシュストレージ (400) に記録されている否かを確認 (s101) する。

【0041】

要求されたデータがキャッシュストレージ (400) にキャッシュされていない場合、要求はメインストレージインタフェース装置 (240) に渡され (s104)、メインストレージ (500) からデータが取り出される。要求が完了すると、処理の終了がホストコンピュータ (100) に通知され (s106)、データを要求するプログラムが残りの処理の完了を待たずに実行を再開することができる。

【0042】

s104 で取り出されたデータが既にキャッシュされている (s107)、或いは、キャッシュストレージインタフェース装置 (250) のデータ転送帯域幅が予め設定した閾値を超過、即ちキャッシュストレージ (400) が過負荷となっている (s108) 場合には、データはキャッシュストレージ (400) にキャッシュされない。

【0043】

s107, s108 いずれも「NO」であれば、つまり取り出されたデータをキャッシュすべき場合には、キャッシュコントローラ (220) がキャッシュディレクトリ (222) を使ってキャッシュストレージ (400) 上の記憶領域を割り当てる (s109)。

【0044】

十分なキャッシュ領域が割り当てられると (s110)、データはキャッシュストレージ (400) に書き込まれ (s112)、キャッシュディレクトリ (222) は書き込み動作の完了を更新する (s113)。キャッシュ領域割当てが失敗した場合、いくつかのキャッシュディレクトリエントリは、キャッシュストレージの領域を空けるために無効化される場合がある (s111)。そのような場合、キャッシュディレクトリエントリの最長時間未使用順を用いるか、あるいはキャッシュヒット率を最高にするのに最適な他のアルゴリズムを用いて無効化するキャッシュディレクトリエントリを選択することができる。

【0045】

要求されたデータがキャッシュストレージ (400) にあると判定された (s101) 場合、キャッシュコントローラ (220) は、ストレージインタフェースモニタ (221) によって監視される負荷と統計情報に応じて、データを取り出すために使用する最適なインタフェース、すなわち、キャッシュストレージインタフェース装置 (250)、メインストレージインタフェース装置 (240) のいずれかを、決定する (s102)。

【0046】

より具体的には、ストレージインタフェースモニタ (221) は、メインストレージインタフェース装置 (240) のデータ転送速度およびキャッシュストレージインタフェース装置 (250) のデータ転送速度を監視するとともに、アクセスされるメインストレージ 500 とキャッシュストレージ 400 とのパフォーマンスの統計データも管理する。ストレージインタフェースモニタ (221) は、上記の監視と管理の結果に基づき、下記の条件判断にしたがって最適なインタフェースを選択する。

(1) キャッシュストレージ (400) のデータ転送速度が最大帯域に近づき、過負荷になった場合には、メインストレージインタフェース装置 (240) を使用する。

(2) メインストレージインタフェース装置 (240) だけを使用することにより、要求された出力の処理時間が短縮されると判断される場合は、メインストレージインタフェース装置 (240) を使用する。具体的には、キャッシュヒットに偏りがある等の理由により、部分的にキャッシュされたデータを無視して要求をメインストレージインタフェース装置 (240) だけを用いて処理する方法が出力要求を二種類のストレージに対する要求に分割する方法より速い、と判断される場合である。

(3) 上記 (1) (2) に当てはまらない場合は、キャッシュストレージインタフェース装置 (250) を使用する。

【0047】

キャッシュストレージインタフェース装置 (250) が選択されると、入出力要求はキャッシュディレクトリ情報に従って変えられ、キャッシュストレージ (400) からデータが取り出される (s105)。メインストレージインタフェース装置 (240) が選択されると、データはメインストレージ (500) から取り出される (s104)。

【0048】

ステップ (s104、s105) で行われた処理は、キャッシュストレージ (400) に部分的にキャッシュされた出力要求参照データに対して並列に行われる。しかし、そうした出力要求は、データの一部がキャッシュストレージに既にキャッシュされていても、ストレージインタフェースモニタ (221) によって提供される情報に応じてステップ 104 で完全に処理することができる。

【0049】

図3は、図1の構成をより詳細に示すもので、デバイスアダプタ (200) として、FC/SCSI デバイスアダプタ (200) の構成を示している。

【0050】

図1のネットワーク (300) とメインストレージ (500) には、SAN (300) と RAID 装置 (500) が相当し、キャッシュストレージ (400) には SCSI ディスク (400) が相当する。図1のメインストレージインタフェース装置 (240) には、SAN (300) を介して RAID 装置 (500) にアクセスするためにファイバチャネル (FC) コントローラ (241) を使用した FC インタフェース装置 (240) が相当する。図1のキャッシュストレージインタフェース装置 2 (250) には SCSI コントローラ (251) を使用した SCSI インタフェース装置 (250) が相当する。

【0051】

FC/SCSI デバイスアダプタ (200) は拡張バス (140) を用いてホストコンピュータ (100) に接続される。キャッシュコントローラ (220) は、ローカルキャッシュの SCSI ディスク (400) の領域割当てを管理するキャッシュディレクトリ (222) を保持する。

【0052】

ストレージインタフェースモニタ (221) は、FC インタフェース装置 (240) のデータ転送速度および SCSI インタフェース装置 (250) のデータ転送速度を監視するとともに、固定の入出力要求のサイズに合わせた入出力要求の平均処理時間の形で、FC および SCSI インタフェースを介してアクセスされる各記憶装置のパフォーマンスの統計データも管理する。

【0053】

ストレージインタフェースモニタ (221) は、上記の監視と管理の結果に基づき、前述と同様の条件判断を行って、最適なインタフェースを選択する。

【0054】

図4に、キャッシュストレージ (500) の領域を固体サイズのブロック単位に分割して管理するキャッシュディレクトリ (222) の例を示す。キャッシュディレクトリエントリはブロック番号 (10) ごとに、ブロックが空いているかあるいはキャッシュされたデータが格納されているかをブロック状態 (11) によって示す。キャッシュされたデータはデバイス識別番号 (デバイスID) (13) で示され、開始セクタ番号 (14) とセクタ数 (15) によって記憶装置上のアドレスとブロック内にキャッシュされたこのアドレスで始まるデータ量が特定される。

【0055】

キャッシュヒットあるいはキャッシュミスは、各キャッシュ入力開始セクタ番号 (14) および対象装置のキャッシュされたデータに格納されているキャッシュデータのブロックのセクタ数 (15) を、SCSI 入出力要求パラメータと比較することによって判断することができる。

【0056】

各キャッシュエントリに、最後にデータにアクセスした時刻を示す最終使用時刻 (12) を付けることにより、キャッシュストレージが一杯、即ちキャッシュブロックが全て使用されている時に、最長時間未使用法を使ってキャッシュデータの置換えを行うことが可能となる。

【0057】

ホストコンピュータ (100) が、図2の s100において、読出し SCSI コマンドを発した時、キャッシュコントローラ (220) はキャッシュエントリを利用してキャッシュされたデータを探索する。キャッシュミスの時には、受け取った SCSI コマンドを FC コントローラ (241) に直接渡して、SAN (300) を介して RAID 装置 (400) からデータを取り出すことができる。データを無事引き出すと、キャッシュコントローラ (220) は、キャッシュディレクトリに従ってキャッシュ領域を割り当ててデータをキャッシュすなわち、SCSI ディスク (400) に書き込むと共に、キャッシュディレクトリ (222) へ、新規にキャッシュされたデータの入力を行う。キャッシュヒット時には、入出力要求完了時間の概算に応じて使用するストレージを選択することによって、最高パフォーマンスを達成し、キャッシュヒット率が高い場合には SAN 帯域幅を他の用途に有効に使用できるようにする。

【0058】

本実施形態のデバイスアダプタ (200) に接続するローカル SCSI ディスク (400) に大容量ディスクを用いることにより、ギガバイトのオーダーのデータをキャッシュすることが可能である。これは、アクセスしたファイルを全てローカルディスク (SCSI ディスク) (400) にキャッシングすることを可能とし、大容量のデータを RAID 装置 (500) から取り出すビデオストリームサーバにとって特に有効である。

【0059】

たとえば、SCSI ディスク (400) のサイズを、ビデオストリームサーバのワーキングセット、即ちビデオストリームサーバを使用するクライアント装置が最も頻繁にアクセスしたファイルサイズの合計を含む全てのビデオファイルの合計、に応じて決定した場合、全ファイルがキャッシュストレージである SCSI ディスク (400) にキャッシュ

され、RAID装置（500）への遠隔アクセスが全て抑制される。これにより、SAN（300）での実質データ転送帯域幅が拡大し、同一のSAN構成でもより多くのストリームサーバを使用可能になることから、システムの拡張性を向上させることができる。

【図面の簡単な説明】

【0060】

【図1】 開示したデバイスアダプタの構成を簡易化して示したブロック図である。

【図2】 開示したデバイスアダプタに読出し入出力要求を発する処理を示した簡易フローチャートである。

【図3】 ファイバーチャネルとSCSIインタフェースを用いた開示したデバイスアダプタの一実施形態の例を簡易化して示したブロック図である。

【図4】 キャッシュディレクトリの実施形態の例である。

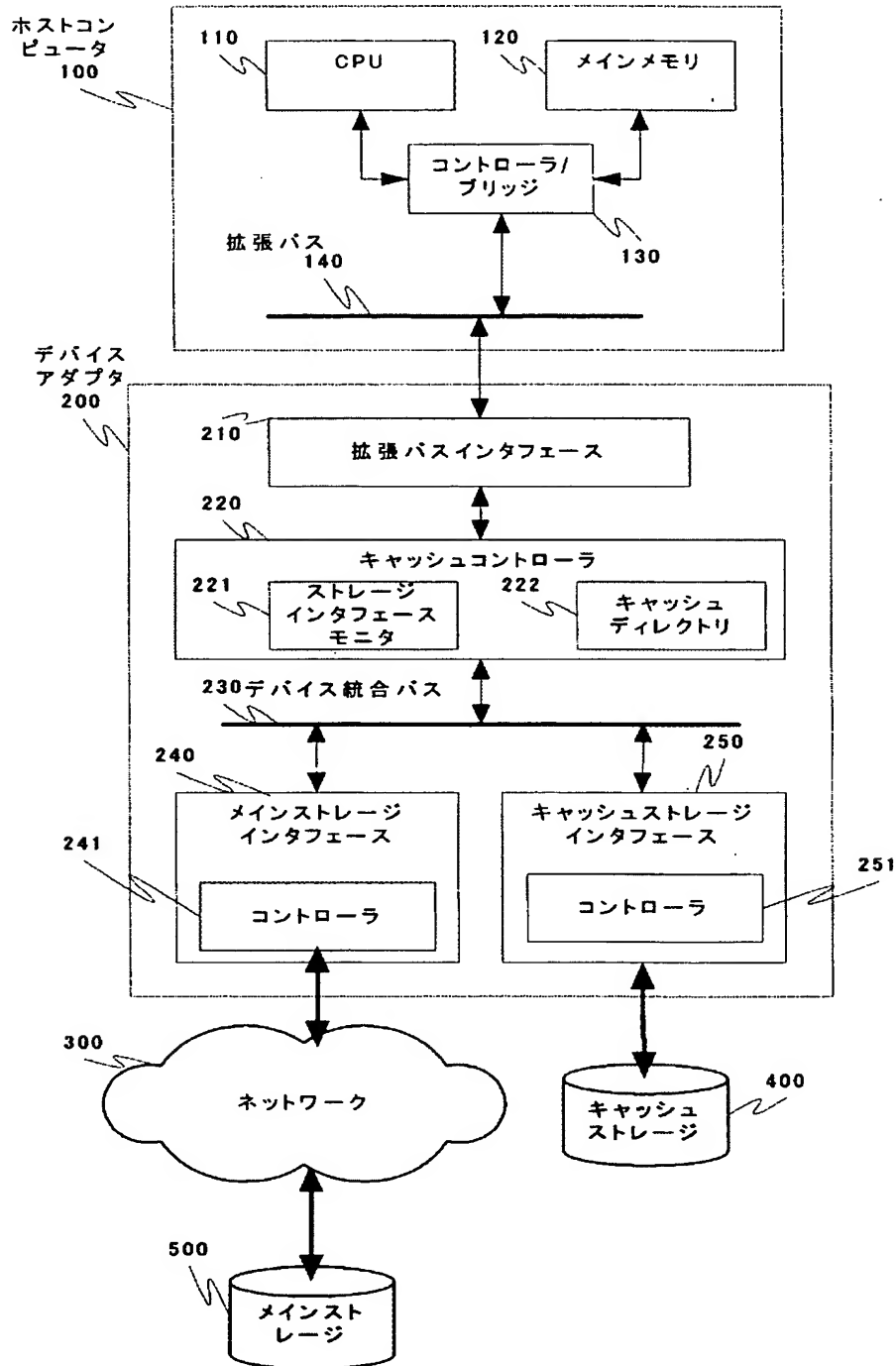
【符号の説明】

【0061】

100…ホストコンピュータ、110…CPU（中央処理装置）、120…メインメモリ、130…コントローラ／ブリッジ、140…拡張バス、200…デバイスアダプタ、210…拡張バスインタフェース装置、220…キャッシュコントローラ、221…ストレージインタフェースモニタ、222…キャッシュディレクトリ、230…デバイス統合バス、240…メインストレージインタフェース装置、241…コントローラ、250…キャッシュストレージインタフェース装置、251…コントローラ、300…ネットワーク、400…キャッシュストレージ、500…メインストレージ。

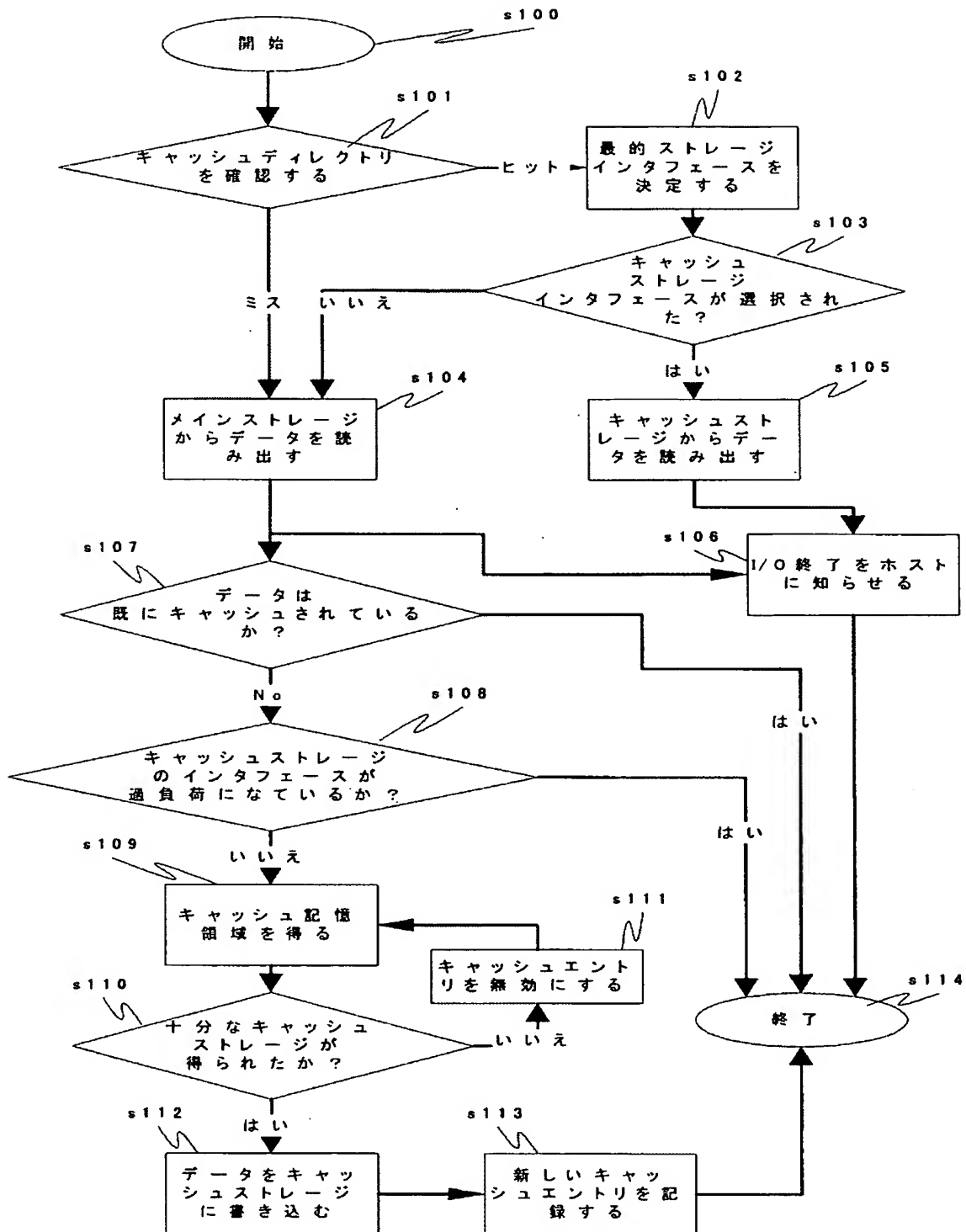
【書類名】 図面
【図 1】

図 1.



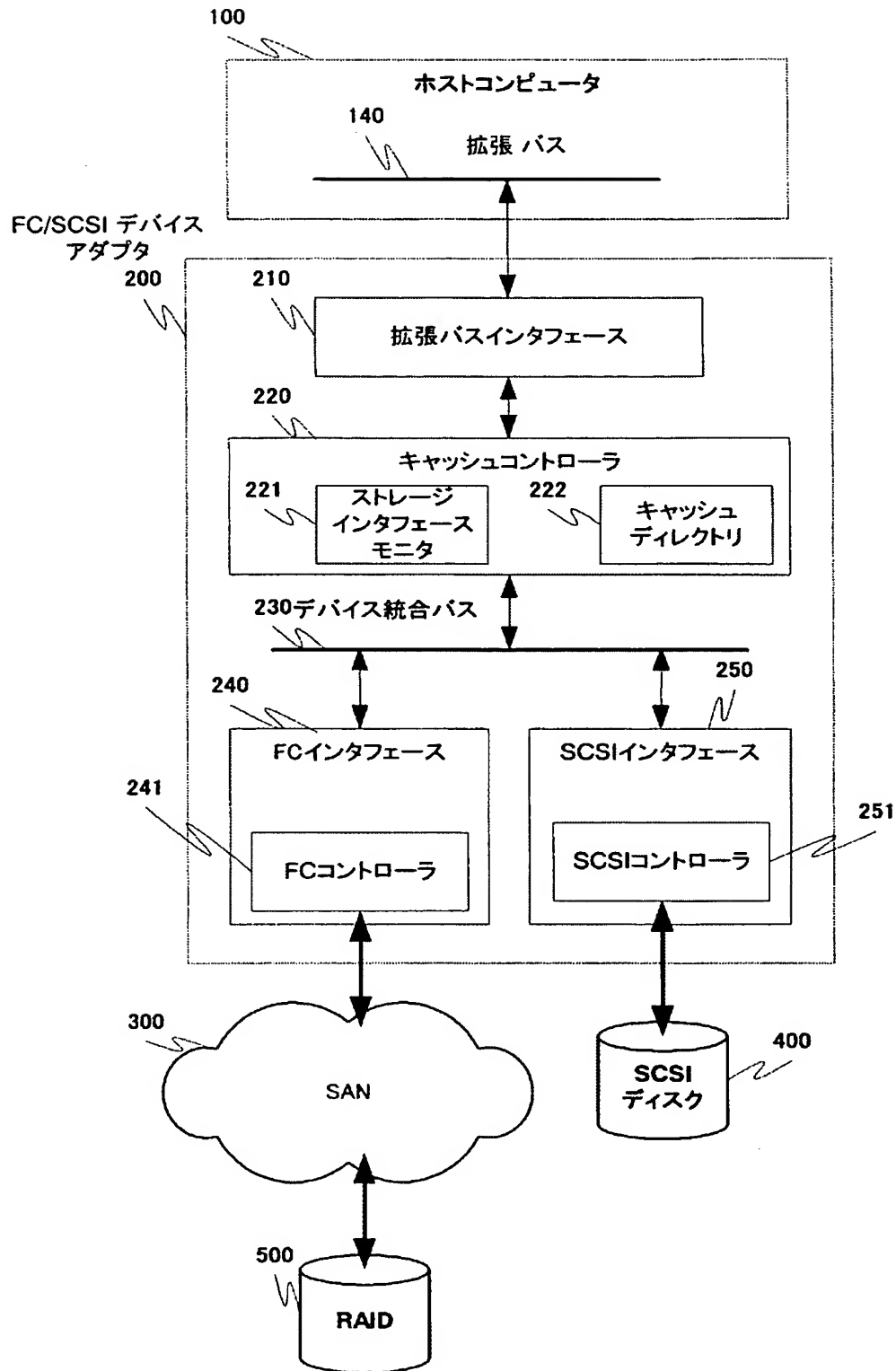
【図 2】

図 2.



【図 3】

図 3.



【図 4】

図 4.

10 h	11 h	12 h	13 h	14 h	15 h
キャッシュブロック 番号	状態	最終使用 時間	デバイスID	開始セクタ	セクタ数
0	使用中	100	1	0	150
1	未使用				
2	使用中	200	2	151	100
⋮					
n	使用中	110	1	1050	27

【書類名】 要約書**【要約】**

【課題】 ストリーミングをクライアントに配信する際にストレージからデータを取り出す能力によってサーバ能力が制限され、データ転送速度が低下し、処理時間が増加する。

【解決手段】 サーバとストレージとを接続するデバイスアダプタは、キャッシュコントローラと、キャッシュ装置にアクセスするキャッシュインタフェースと、ストレージにアクセスするストレージインタフェースとを備える。キャッシュコントローラによって、ホストコンピュータからは意識されずに、ストレージから取り出したデータをキャッシュに格納できる。キャッシュコントローラは、キャッシュ上の記憶領域割当を管理するキャッシュディレクトリと、ストレージとキャッシュの負荷とパフォーマンスの情報を維持するモニタとを備える。ストレージとキャッシュは、キャッシュ内の要求データの存在だけでなく、ストレージの現在の負荷とパフォーマンスによっても使い分ける。

【選択図】 図 2

認定・付加情報

特許出願の番号	特願 2 0 0 3 - 3 8 3 2 4 8
受付番号	5 0 3 0 1 8 7 4 7 8 7
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 5 年 1 1 月 1 4 日

< 認定情報・付加情報 >

【提出日】 平成 15 年 11 月 13 日

特願 2 0 0 3 - 3 8 3 2 4 8

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1 . 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所